Washback Effect of EFL Teacher-Made Test on Teaching-Learning Process of Communicative English Skills Course at Ambo University

Metages Gebeyehu Alebachew

Lecturer

Department of English Language and Literature

Fiche University, Oromia, Ethiopia

Motuma Hirpassa Minda

Assistant Professor

Department of English Language and Literature

Ambo University, Oromia, Ethiopia

maldamercy@yahoo.com

Abstract

Prompted by the problem the effects of washback, this study was mainly intended to evaluate the washback effect of EFL teacher-made tests/examination on teaching-learning process of Communicative English Skills at Ambo University. Specifically, it examined the influence of the examination on the domains of teaching and learning. It further explored mediating factors of wash back on teachers and students' practices. To this end, the data were collected from three years (2008, 2009, 2010 E.C.) communicative English skills course teachers' made examinations through document analysis method. Supplemental data were also collected from EFL teachers and English major

students respectively using interview and questionnaire. Both teachers and the students were selected on the bases of randomly sapling techniques. The study employed mixed methods to analyze and interpret the results of the study. The findings of the study reveal that teacher-made examination affected the domains of teaching-learning process negatively. This means it affects students' learning attitudes, motivations, and readiness in leaning some components of the materials, and experience in taking examination, practice in using the learning materials and methods in general. Some components of the course, such as writing and listening, were not included in the final examination of the course. As a result, students were influenced to give more time for the work of previous examination questions rather than studying the materials. Based on the findings of the study, teachers are recommended to set examination questions from the contents of the materials they used to teach the components of the course and give equal emphasis for all the components of the language during teaching-learning process and assessment. As the committee members usually set the examinations, the teaching material choice should be amended to align the contents of the examination questions with the contents in the teaching materials.

Keywords: Backwash, Alignment, Content Validity, Reliability, Authenticity.

Background of the Study

Language teaching and testing cannot appear separately. Various scholars attempt to define the relationships of teaching and testing language. Thaidan (2015), Djuric (2008) argue that the process of teaching a foreign language usually involves three major steps: the teaching acts of presenting and explaining new material, providing practice and testing. The term "test" is generally interpreted as a mode of assessing a learner's ability, knowledge, or performance in a given subject area (Brown, 2004; Iyer, 2015). Relevant, comprehensive and valid tests measure learners' language ability and performance (Ching, 2008; Zucker, 2005). The impact or effect of tests on teaching and

learning process is referred to as *systemic validity* (Creswell, 2009), *test impact* (Cohen, et.al. 2007; Brown, 2012), *consequential validity* (Anderson and Bachman, 2008), *backwash* (Hughes, 2003), and *washback* (Brow, 2004). The aim of all the terms is improving teaching and learning in language. For consistency, 'washback' is used in this study. In this background section, four concepts of washback were reviewed: 1) What is washback? 2) How does washback work? 3) How can we promote positive washback? 4) How can we investigate washback?

Several scholars have attempted to define washback. The term 'backwash/washback' has been used to refer to the technique a test affects teaching–learning process (Weir, 2005; Hughes, 2003). Brown (2004) argues that washback is generally refers to the effects tests have on instruction in terms of how students are ready for the test. Washback of test is dependent on the mismatch or proper match of the course with curriculum contents and types of materials, teachers activities, the objectives of the course and test, students' feeling and motivation towards learning the language and teaching methodology and testing techniques (Hughes, 2003; Thaidan, 2015; Zucker, 2005). According to Burger (2008), washback effect is not simply good or bad. Teaching-learning practice that might occur with or without the test is rather a good or bad practice that is linked to the effect or impact of test (Green, 2007). In this perspective, language tests have either positive or negative effect on the language teaching and learning. This means test with a good quality (i.e., high content validity, construct validity, reliability and practicality) may have a positive/beneficial impact/effect on language teaching leaning processes, and the reverse is also true that unexpected side effects of tests are negative/harmful (Khan, 2013). In both consequences of test, washback effect has both dependent (washback effect) and independent (test) (Hughes 2003).

Similarly, Hughes (2003); Thaidan (2015) also define the washback effect as the outcome of a test or an examination which results either in positive or in a negative way. Positive washback effect

occurs when there is harmony between the objectives of the teaching materials and the students' examination results or a class test performance (Gumaa, 2010) and Sprtta, 2005). The negative washback effect occurs when there is no harmonization between what is taught and what is performed (such as narrowing down the content) Wier, 2005). Both of these types of washback influence the teaching as well as the learning process (Zuker, 2005). Whatsoever, the definition is, washback impacts or effects have long been believed that tests directly influence educational processes in various ways (Taylor 2005).

In this study, building on the 'Washback Hypothesis' proposed by Anderson and Bachman (2008) and suggestions from Hughes (2003), washback works with a model that identifies participants, processes and products which may influence, or be influenced by effects of teacher-made test. According to the 'Washback Hypothesis' proposition, test influence learning, what learners learn, how learners learn, the rate and sequence of learning, the degree and depth of learning, the attitudes to the content, method, etc. of teaching and learning, the attitudes and perceptions of the participants towards their teaching and learning tasks (Anderson and Bachman, 2008). These may in turn influence what the participants do in carrying out their work (process). According to Hughes (2003), such processes include materials development, syllabus design, changes in teaching methods or content, learning and/or test-taking strategies, etc. The following figure 1 shows the basic model of washback.

Figure 1: A basic model of washback (Source: Iyer (2015)

The above model of washback implies that practitioners should have a clear understanding how test washback effects affect the participants (teachers, students, researchers, material writers and curriculum designers), teaching learning process and products (learning, teaching, new material and research result) (Iyer, 2015). According to Iyer (2015), the areas of washback are attitudes, behaviour, and motivation of teachers, learners and parents, wider social feelings, classroom pedagogy, curriculum development, and educational systems and policy, administrators, content of the language lessons (Sukyadi and Mardiani, 2011; Taylor, 2005).

Scholars have also forwarded different techniques of promoting positive washback effects. According to Iyer (2015), testers want to promote positive washback effects of tests; they need to test the abilities and developments they want to encourage, sample widely and unpredictably, use direct testing, make test criterion-referenced, base on achievement of objectives, ensure that test is known and understood by students and teachers when necessary, provide assistance to teachers. Newfield (2005) stressed on the timely and detailed provision of test scores that are easily interpretable to foster positive washback effect from testing. Test scores should be credible for the test-takers. In addition,

Taylor (2005) pointed out that tests should incorporate authentic tasks that lead to achieve language-learning goals and to foster learner autonomy and self-assessment. The potential backwash effect should join validity and reliability in the balance against practicality (Hughes, 2003). In this case, Sukyadi and Mardiani, (2011) assert that the concept of the washback effect of the test includes validity, reliability, authenticity and practicality the test.

To find out whether testing will bring positive washback effect on teaching and learning or not, Taylor (2005) posed the following eight questions:

> *...do the participants understand the purpose(s) of the test and the intended use(s) of the results? Are the results provided in a clear, informative and timely fashion? Are the results perceived as believable and fair by the participants? Does the test measure what the program intends to teach? Is the test based on clearly clear goals and objectives? Is the test based on sound theoretical principles which have current credibility in the field? Does the test utilize authentic texts and authentic tasks? Are the participants invested in the assessment activity? (pp. 276-277)*

The concept of examination washback effects can be studied in terms of two dimensions: micro effects and/or macro effects (Cheng, et al., 2004, Hughes, 2003; Thaidan, 2015). Some language testers describe effect/outcome as micro washback and consider impact as macro washback in the language context (Taylor, 2005). In this regard, (Pan, 2009) states that the effect of test on classroom teaching and learning is micro washback whereas the impact of test on the community as a whole is referred to as test macro washback scenarios (Sadeghi, 2014; Thaidan, 2015). However, most testers locate both concepts within the theoretical notion of 'consequential validity' in which the social consequences of testing are part of a broader, unified concept of test validity (Messick, 1996;

Bailey, 1996). The following diagram 2 specifies in the components of washback effects: macro and
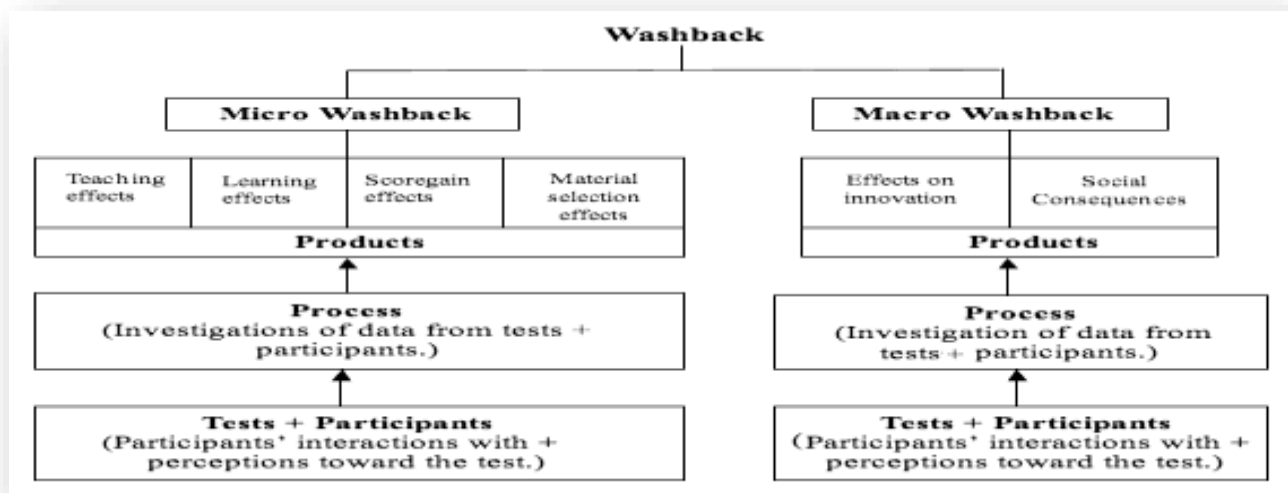
Micro levels.

Figure 2: Scope of washback (Source: Thaidan, 2015)

Source: Thaidan (2015)

The above figure reveals the components of washback effects. The first is macro washback

effects. Macro washback effects refer to the influences of tests in which it influences a society and are

influenced by the social climates. Studies on macro washback effects can be conducted on high-stakes

tests, which incorporate a large number of test-takers and major decisions making (Thaidan, 2015;

Messick, 1996; Bailey, 1996). In this case, studies focus on the effects of tests and test results on the

professional and employment opportunities, teaching learning system and society more widely

(Thaidan, 2015). The second part of washback is micro washback. Micro-washback effects refer to

low-stakes tests, such as in-class tests, teacher-made tests, classroom assessments or teachers'

assessments (Thaidan, 2015; Sadeghi, 2014).

This study was therefore limited to micro level or teacher-made test washback effects.

Teacher-made/micro washback refers to an assessment carried out by teachers in the classroom as a

central part of the educational process (Anderson and Bachman, 2006; Spratt, 2005). The purpose of

teacher-made test /micro washback is to describe instructional purposes, influence what students study, and help instructors to obtain perspective on their courses (Thaidan, 2015). Teacher-made tests are normally prepared and administered for testing classroom achievement of learners, evaluating the method of teaching adopted by the teacher (Madsen, 1983). Teacher made-test is designed to solve the problem or check the requirements of the class for which it is prepared (Djuric, 2008; Spratt, 2005). In short, micro washback/teacher-made tests are the most worthy instrument in the hands of teacher (Shih, 2007). However, micro washback has limitation cover a broad range of abilities and skills in language skills (Green, 2007). Teacher-made tests measure a limited part of a subject area in a very helpful way to define the short-term course objectives envisioned by the teacher (Thaidan, 2015). However, the validity of teacher-made examination remains debatable in the educational assessment process.

Previously, various studies have globally and locally been conducted in different areas of test washback effects. Globally, Djuric (2008); Green (2007); Shih (2007); Thaidan (2015) have studied on English National Examination washback effects and found out the Situations of positive and negative washback effects. Locally, Simachew (2012) and Yasin (2014) have respectively studied the washback effects of University Entrance Examination and Ethiopian general secondary education certificate in English examinations. Hailu (2015) have also assessed the content validity of EFL teacher-made tests of grade 11 at Ambo preparatory school. Motuma (2017) has studied the alignment of grade eight national examinations to its textbooks. However, no study has been conducted on university EFL teacher-made tests/examinations washback effects on teaching-learning process particularly on the communicative English skills course that ha complex components of language objectives. For this reason, this article evaluates the washback effects of EFL teacher–made tests on the teaching-learning processes of communicative English skills course (Spratt, 2005). Specifically, the current article has examined how teacher-made tests influence the teachers and students'

perception, attitude and motivation, activities and selection of teaching-learning materials in the course.

Statement of the problem

The core of the problem is concerned with the assumption that the EFL teacher-made tests fail to measure the entire components of the Communicative English course contents (Hawkey, 2006; Ambachew, 2003). Teacher-made tests normally measure only a limited part of abilities and rely too heavily on memorizing facts and procedures (Hawkey, 2006; Hughes, 2003). Teacher-made test washback effect is debatable because it is less valid and reliable against practicality (Motuma, 2014; Hughes, 2003). This means, the result of the communicative English skills course teachers-made test usually create grievance because of its inconsistent for most students (Nugussie, 2002; Motuma, 2015). On the other hand, most (67%) of Ambo University students are trying to cheat at least on one or two sort of assessment tools the teachers use (Motuma, 2015). These students are not conscious about the effect of tests and test results on their future professional career (For example, access to educational and employment opportunities, Taylor, 2005; Nugussie, 2002).

Scholars believe that university teacher-made tests are the base for teaching learning process and for the students' future career. Teachers-made tests determine the classroom activities of both teachers and students (Davies, 2014; Nugussie, 2002). However, with the expansion of universities, the problem of teacher-made washback effects has become more and more apparent at the level. Though universities have almost harmonized curriculum and similar directives, policies, rules or regulations of assessments and examination procedure set by MOE, the teachers and the students have been violating them. Yet, knowing this severe and pressing concern, no investigate has been conducted to examine the teacher-made test washback effects (Motuma, 2017). This implies that teacher-made tests results have been given less emphasis particularly in the Ethiopian universities language studies. Thus, this study has intended to address the following research questions.

1. How do teacher-made examinations of the course affect teachers' teaching practices in Communicative English Skills Course?

2. How do teacher-made examinations of the course affect students' learning activities in communicative English skills course?

3. How do communicative English skills course teacher-made examinations influence teachers and students' material selections for communicative English skills course teaching learning?

4. To what extent do teacher-made examinations influence the teaching learning processes of the course?

Methodology of the study

In the view of pragmatic paradigm and mixed research design, a descriptive survey method was employed in this study to evaluate the washback effects of EFL teacher–made tests on the teaching-learning processes of communicative English skills course (Kothari, 2004; Kumar, 2011). This method is used to answer the questions as well as to assist a researcher to address the research problem. Besides, mixed research approach helps to provide both quantitative and qualitative answers to the questions of "who", "what", "why", and "how". Based on this pragmatic research paradigm, the population, sample size and sampling techniques, data gathering tools and data analysis methods were presented below.

The major sources of the data were communicative English Skills Course teaching materials (modules), and examinations, first-year and second-year English major students who were registered to communicative English skills course in 2009-2010 and full-time English language teachers who teaching the course in 2008-2010 at Ambo Universities. Eight-eight (44 first-years and 44 second-years) students were enrolled respectively in 2009 and 2010 to the department of English language and literature. Three of the 42 EFL teachers were teaching the course for the English major

students during the time of study. However, a committee with five members prepared the examinations each year. Thus, all the 88 students were included based on the available sampling techniques, and 12 EFL teachers (those who taught the course and who set the examinations) were included to the study based on purposive sampling method.

To triangulate the data gathering instruments to gauge micro levels of washback: document analysis, pre-and-post tests interview with students, semi-structured interview with teachers, and questionnaires (Ching, 2008) were used. Document Analysis is focused on the analyses of the contents of examinations and teaching module (Davies, 2014). The aim of content analyses was to understand the alignment/relationship of the course module and the tests contents and objectives. In addition to document analyses, based on the recommendations of Ching (2008) to gauge micro levels of washback, pre-and-post tests interview was conducted with nine students (three students with top, three students with medium and three students with lower scores in their previous examination results). Next, **a** questionnaire with open-ended and close-ended questions was used to gather the opinion of first and second year students on the relationship of the examinations and the learning module. Finally, semi-structured interview was also conducted with the teachers who were teaching course and sat the examinations (Shannon-Baker (2016).

Methods of Data Analysis

In this study, both quantitative and qualitative data analyses methods were employed to describe the micro washback effects of teacher-made examinations on teaching-learning process of communicative English skills course. Using integrative data analysis approach, qualitative information was described verbally. To achieve the quantitative data analysis, descriptive statistics (frequency percentage, mean, standard deviation) and Pearson Sidney Siegel's formula

(contingency coefficient 'C') are used to make judgments about the nature of washback effects of the teacher-made examinations. The contents analysis followed three steps. The first step was the analysis of the contents/activities (to determine the number or frequency of contents) of the communicative English skills course materials, which was organized in to six major components: Speaking, Reading, Vocabulary, Grammar, Listening, Writing. The second was analyzing the pre- and post- test interview with students. The next step is to determine the frequency of contents of the items in each examination in a similar category.   The third step was the computation of the Siegel's contingency coefficient (C) to determine the level of relationship (Washback effect) of the examination with the course materials.To calculate the Siegel's contingency coefficient 'C', the following formula and interpretations were used (Kothari, 2004).

$$C = \sqrt{\frac{x^2}{N + x^2}}$$

Where, C= the value of relationship X^2= refers to Chi-square

   N=grand total

     According   to   Sidney   Siegel's   formula   (Green,   20007),   contingency   coefficient (relationship/correlation) between two observations are interpreted as in the following table 1.

No.	Value of contingency		
1	C= 1	Perfect correlation, comprehensiveness	Positive
2	C= 0.91-0.99	Very high correlation, comprehensiveness	Positive
3	C= 0.71-0.91	High correlation, comprehensiveness	Positive
4	C= 0.41-0.70	Moderate correlation, comprehensives	Positive
5	C= 0.21-0.40	Low correlation, comprehensives	Negative
6	C= > 0.2	Very low correlation, comprehensives	Negative

Table 1: Interpretation of the washback effect (Source: Green, 20007)

Results and Discussions

Results

        As it has already been indicated in the methodology section, the integrative analysis of the data pursues three steps: the analysis course materials, determination of the frequency examinations and the computation of the contingency coefficient and washback effect. To determine the level of the relationship between the examination the course, the Siegel's contingency coefficients were calculate for each case, and the result was interpreted accordingly. Finally, the result was summarized and discussed to make it as clear as possible for the reader. Therefore, the results and discussions of the document analyses, interviews and questionnaire were presented in integrative approach. The following table the frequency of the contents of the course materials from 2008 to 2010 E.C.

The Result of Course materials' Analysis

        Course material analysis is the main source of data to determine the alignment of the course with the tests/examinations. Thus, the communicative English Skills Course is usually given to first-year students at first semester. The teaching material/module of the course is usually revised every year. The module commonly has five units which of each unit have sex major language content areas or components: reading, vocabulary, grammar, speaking, listening and writing. The sequence of the components/contents of the course marital varies from chapter to chapter within a module. The following table shows the coverage of language components in the course material. The findings in the table summarized from analysis of teaching materials (document analyses).

No	Content area	Frequency in each Years			Total frequency	%	Rank
		2008	2009	2010			
1	Speaking	27	53	27	107	10 %	5
2	Reading	119	121	119	359	33.4 %	1
3	Vocabulary	57	86	57	200	18.6 %	2
4	Grammar	48	86	48	182	17 %	3
5	Listening	59	62	59	180	16.8 %	4
6	Writing	18	9	18	45	4.2 %	6
	Total	**328**	**417**	**328**	**1073**	**100 %**	

Source: *Communicative English Skills Course Modules*

Table 2*:* Frequency of the contents of the course materials from 2008 to 2010 E.C

The table shows the frequency of contents of the components of the course in each year, the total frequency of the three years and the frequency percentages of the total contents of the teaching materials. Total content of the course teaching material is 1073 in the years 2008 to 2010. The maximum number (417) of frequency of contents of the Communicative English skills course was observed in 2009 E.C. Among the components of the course, reading activities took the lion's share, 359 (33.4%) that show the modules were dominated by reading activities. Writing was the list major language component in the modules throughout the years, 18 (5.5%), 09(2.2%) and 18(5.5%) respectively in 2008, 2009 and 2010 E.C as well as total 45(4.2%). The distribution of language contents in the course teaching material in descending order are reading (33.4%), vocabulary (18.6 %), grammar (17 %), listening (16.8 %), speaking (10 %) and writing (4.2%). The following section analyses of the frequency of the contents of the examinations given during the years 2008 to 2010 E.C.

Analyses of the Examinations

The analyses of the contents in the components of the course material were used as the basis for the analysis of contents in the examination questions because the contents of each item of the examinations need to be classified according to the major language content areas in the modules. To achieve the representativeness of the examinations, the researcher counted all the items of the examinations given during 2006 to 2009 E. C. and sorted them in to the six major components of the course. The table 3 shows that a total of 138 items of communicative English skills examinations were set during 2008 to 2010 E.C. Specifically, the maximum number of items (**50)** were set in the 2009 E.C. examination as well as 48 and 40 items were set respectively during 2010 and 2008 E.C. examinations.

The comprehensiveness and proportionality of the items of the examinations to the six major language content areas in terms of periods allotted to them. The content of all the communicative English skills course examinations contained only the major language contents: reading (34%), grammar (31%), vocabulary (20%) and speaking (14.5%) to measure the students' abilities and performances. However, listening and writing were ignored from the examinations throughout the years. However, this does not mean that students' performances in writing and listening were not assessed. The EFL teachers confirmed during the interview time that though these two skills were ignored from the final examinations, they were assessed using continuous assessment techniques. The students replied to the questionnaire that they were being assesses in these skills during the continuous assessment.

Table 3: Frequency of the contents of the items in the examinations in 2008 to 2010 E.C

No	Content area	frequency in each Years			Total frequency	%	Rank
		2008	2009	2010			
1	Speaking	6	8	6	20	14.5 %	4
2	Reading	15	12	20	47	34 %	1
3	Vocabulary	4	14	10	28	20.3 %	3
4	Grammar	15	16	12	43	31.2 %	2
5	Listening	-	-	-	-	-	5
6	Writing	-	-	-	-	-	6
Total		40	50	48	138	100%	
Source: *Examination Buckets of the years 2008-2010*							

The table presents the comparisons examinations' items among the language components and the time of examinations. Thereby the distribution of items in the examinations shows that vocabulary questions were highly fluctuated from year to year. Vocabulary questions were 4 in 2008, 14 and 10 items respectively in 2009 and 2010 E.C. Reading has a big (41.7%) coverage in the examination of 2010 E.C. From the 50 examination items in 2009 E.C., grammar covers the highest percentage 16 (32%). The following section presents the comparisons/relationships between the frequency percentages of the contents of the teaching materials and the teacher-made examination in general in the communicative English skills course during the three years.

Relationships between Textbooks and Examinations

There were some observable mismatch representation and disproportion between the contents in the course materials and examinations. In the materials of course, reading, vocabulary, grammar, listening, speaking and writing constituted in descending order covered respectively 33%, 20%, 17%, 17%, 10% and 4%. This shows that the language components received a varying degree of attention because of the objectives of the course. The following diagram three (pie chart) shows the comparison between the frequency percentages of the contents of the teaching materials and the teacher-made examination in general in the communicative English skills course during the three years.

Source: Teaching Modules and Examinations' Bucklets of the Years 2008-2010 E.C.
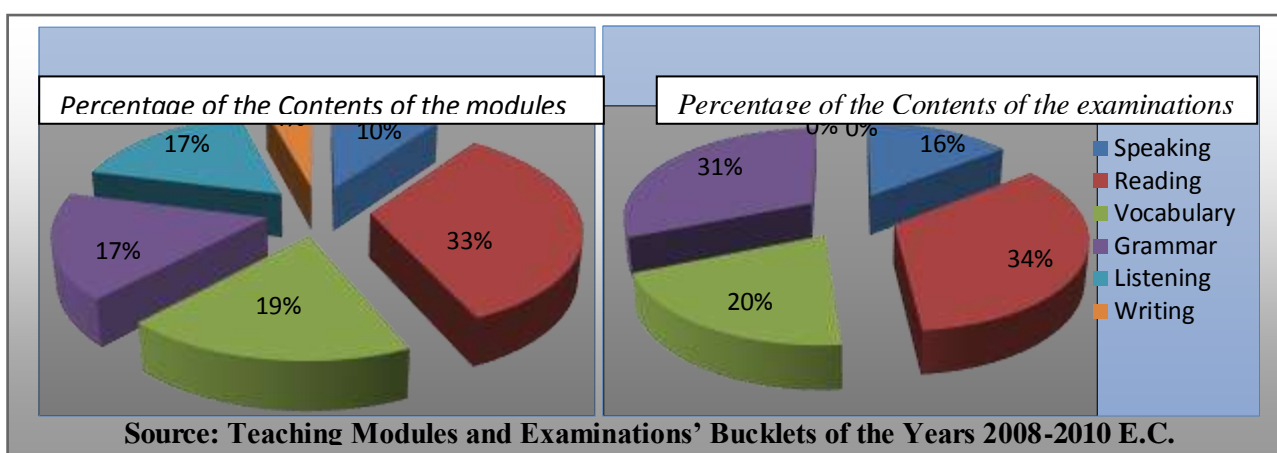
Figure 3: Comparison of Contents in the Course Material and Examinations

The figure 3 relates the contents of the language components of the course module to the contents in the examinations during the three years. Some language components such as reading (33%) and vocabulary (19%) in the course module received similar and higher proportions, respectively 34% and 20% for reading and vocabulary. However, the contents in grammar (31%), listening and writing (0%) in examination do not seem to represent the teaching modules. This is because grammar, listening and writing respectively constitute (31%), 17% and (4%) in the Communicative English Skills Course materials. This indicates that the distribution of language contents in the sampled examination was disproportional to the distribution of the language contents in

the course materials. Such disproportional distributions of contents were the indicators of negative washback effect on teaching-learning process in the course. The following figure 4 also discusses the relationships of the contents of the course materials, contents in the examinations items and the time allotted to the components of the language in the course.

It was difficult to determine exactly the frequencies of time for the contents in the course materials separately because the contents of some of the exercises/skills were either too general or overlapping with each other. However, attempts were made to count the contents and the periods allotted to them as closer as possible to the reality and categorized them based on the judgement of the researcher. The relationships among the contents of the frequencies obtained from the analysis of the course materials, the items of examinations and the time allotted to cover the course materials have been presented in Figure 4 below. The Relationship of the contents of the examinations, module and time by frequency percentages.

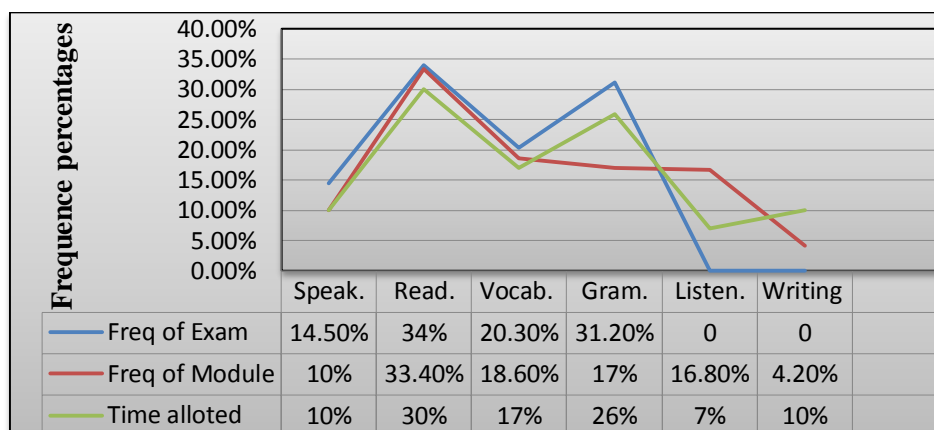|	Speak.	Read.	Vocab.	Gram.	Listen.	Writing
Freq of Exam	14.50%	34%	20.30%	31.20%	0	0
Freq of Module	10%	33.40%	18.60%	17%	16.80%	4.20%
Time alloted	10%	30%	17%	26%	7%	10%

Figure 3: The Relationship of the Contents of the Examinations Module and Time by Percentage

The figure shows close estimation of periods to cover the language contents in the teaching materials, but the contents in the items of the examinations were not comprehensive and proportional to the time and the contents in the teaching materials. For example, 10% and 07% of the time was allotted

respectively to the writing and listening sections, whereas no question from listening and writing was included in examination. The next section discusses the two sets of data (contents of teaching materials and the examinations items) to determine the degree of washback of the examination as it has already indicated in methodology section. The total frequency of the major contents in the teaching materials and examinations of the communicative English Skills course are summarized in table 4. Finally, based on the summary of the data, Sidney Siegel's Correlation Coefficients were computed.

Table 4: Summary of Sidney Siegel's Correlation Coefficients and Washback Effects

S/N	Years in E.C.	Relationship By Years			Washback Effects	Remarks
		N	**X^2**	**C**		
1	2008	**376**	63.549	0.38	*Negative*	*Low correlation*
2	2009	**467**	12.5123	0.158	*Negative*	*Very low*
3	2010	**376**	**30.0962**	0. 273	*Negative*	*Low correlation*
4	Total	**1211**	**65.516**	0.115	*Negative*	*Very low correlation*

Source: Computed from the previous tables and figures

The data in Table 4 presents the sum of the total expected value/frequencies (1211) of the teaching materials, which was computed from the total observable value (1073) with 138 observable values in the examination. The overall result of Sidney Siegel's contingency coefficient *(C = 0.115)* reveals that the alignment between the examinations given during 2008-2010 E.C and the teaching materials of the course is very low. The table indicates a clear disproportion between the content of the teaching materials and examinations in general. The information in table 4 reveals the degree of the relationships between the contents of the teaching materials and examinations' items separately by year. The degree of the extent of relationship between the contents of the materials and the examination found to be C = 0.38, C =0.1158 and C = 0.273 respectively for the years 2008, 2009, 2010. These Sidney Siegel's Correlation Coefficient values reveal that the examinations of

communicative English Skills have negative/harmful washback effects on the teaching–learning processes of the course during the sample years.

Discussion

The discussion of the results integrated the data obtained from documents' analysis, teachers' interview, students' pre- and post- tests interview and students' questionnaire. Descriptive statistics such as frequency percentages were used to compare the contents of materials and teacher-made examinations of communicative English Skills course. Using Siegel Contingency Coefficient "C", the extent of relationship between the contents of course materials, estimated time allotted and teacher-made examinations has been determined. The level of relationship between these parameters shows the washback effects of the examination (Hughes, 2003; Cheng et al., 2004). Compared to the teaching materials, reading, grammar, vocabulary and speaking items dominated the examinations. These discussions and of the results were based on the findings summarized in Tables 2-4 and Figures 3 and 4.

The study aimed at evaluation of the washback effect of EFL teacher-made tests on teaching-learning process of communicative English skills course at Ambo University. Objectives of course materials are expected to be similar with the objectives of the examinations. The general objectives of the course state that upon completing the course, students would be able to express their idea in various communicative context, use various vocabulary learning strategies and techniques, write and present reports, read various materials and make their own notes, identify the structure of oral and written discourse (Corse Module, 2005). The specific objectives of every unit expect students to use different language structures and expressions, introduce themselves and others, read texts and answer questions using different reading techniques. Students are also supposed to workout meaning of words from context to learn vocabulary for communication, analyze grammar in real texts to see

how changing the grammar changes the meaning, write descriptive paragraphs to describe themselves, listen actively and take notes in a particular occasion. The alignment of the objectives of the course with the objectives of the examinations has a positive effect on the teaching–learning processes of the course (Hugest, 2003). In other words, the items in the examinations should reflect the objectives of the course.

Examination in communicative English skills course is fundamentally different from other language course (Nigussie, 2002). The communicative English skills course has complex components of domains of objectives to be tested using different type of items that (Burger, 2008; Wall 2000). The objectives of the course were designed to develop students' communicative ability, reading, speaking, listening, and writing skills including the sub-skills: grammar and vocabulary. Students were supposed to practice the intended language skills accordingly to be successful both in their academic development and for their career progress in the world of work. However, the examination tasks did not incorporate all language skills in the course, evidently, writing and listening skills items were missed from the examinations. The overall intentions of the courses were to promote all language skills ability of the students through involving them in a variety of activities as stipulated in the objectives of the course materials.

The results of the study reveal almost negative washback effect of the teacher-Made examination on the teaching learning processes of communicative English Skills course during the sample years. The degree of the extent of cumulative relationship between the content of the teaching materials and the teacher–made examinations found to be very low (0.115). This generally shows that the examinations have negative effect on teaching and learning process of the course. Specifically, the Correlation Coefficient values "C" is 0.38 for 2008, 0.158 for 2009 and 0.273 for 2010 E.C. were evidences for the teacher-made examinations' negative effects on the processes of teaching the course.

According to Hughes (2003) and Cheng et al. (2004), the mismatch of the content of the examinations and the classroom textbook has a negative consequence on teaching learning process.

Comparing with the teaching materials and course breakdown, every year, the booklet of the examination was expected to include all the major language contents like speaking, reading, listening, writing, grammar and vocabulary. The total contents of the communicative English skills course examination consisted of 40, 50 and 48 items respectively in 2008, 2009 and 2010 E. C. However, the examinations of the years were constructed without writing and listening skills. The highest portion (65.2%) of the examinations was occupied by reading (34) and grammar (31.2%). The other language contents have been given less attention; others were neglected in communicative English skills course during examination 2008, 2009 and 2010 E.C. The results therefore confirm the argument that the distributions of communicative English skills course examinations items were comprehensive and disproportional to the teaching materials in the three years, which means the examinations have negative washback effects on the teaching-learning of the course (Iyer, 2015).

Communicative English language teachers asserted that listening skill was mostly neglected in the final examination because of the problems of technological devices. Both teachers and students consider teaching-learning process of listening skills in classroom is challenging. As a result, the examination committee members do not usually incorporate listening skill items in the final in communicative English course examination.This is because there is sufficient of number of computers to test listening skill in laboratory rooms. For these reasons, most teachers and students were ignoring listening skill in the teaching learning process of the course. However, the students' listening skills is assessed during the continuous assessment time. As most of teachers usually skip over the practice of listening skill, both teachers and students do not expect items from listening skills in the final examinations. Regarding this, Hughes (2003) stated that skipping over some portions of course

content can create a negative washback effect. This shows that examination's Washback effect influence what teachers teach, how teachers teach and the rate and sequence of teaching, the degree and depth of teaching, attitudes to the content, method, etc. of teaching and learning (Hughes, 2003). Therefore, communicative English skills final examinations influenced teachers to skip the practice of listening skill activities in the course content. Shih (2007) argues that students and teachers' perceptions of tests affect their learning practices.

The responses of the students to the questionnaire revealed that most students have experienced the influence of the examination on their learning. Normally, students believe that continuous, comprehensive, and proportional examination items to what they have learnt encourage them to work hard, use the teaching materials and time properly. This is because washback effects of examinations influence students' perception, attitude and motivation towards the skills they are learning. The contents of the questions in the final examination change the students' perceptions to use reference materials, time and methods of learning and studying the entire course contents. In general, the course examination has harmful washbach effects on learners/students that they fear, ignore and late to class. In relation to writing, correct punctuation and spelling, sentence, paragraph, completing dialogue, technical aspects of writing such as letter types and formats of writing were included in the teaching materials. Hence, comparing with other skills, students were not equally motivated for listening and writing skills. This confirms that the consequence of the examination was negative washback effect.

Finally, to distinguish the EFL teachers and students' choices, the kind of teaching materials they frequently use and why they preferred these materials were analyzed.The major question asked in this sub-topic was: "Do communicative English Skills course teachers use additional materials in teaching the course? What kinds of reference materials do they usually use? Why do they prefer the materials the usually use?" Most of the teachers and students try to use additional materials to align the

teaching learning processes to the nature of the final examinations of the course. The purpose of the selection of the materials is to help students be ready for the final examination. This reveals that course examinations influenced the teachers' choice of materials. All the chosen materials were grammar rules, vocabulary and other language skills. Learners' responses also show that the materials (books) usually related to read compressions, grammar, vocabulary and writing books that are related to the previous examination. This shows the harmful washback effects of the final examination on teachers and students materials preferences. They search for examination related materials to meet their expectations of examination content overages.

Conclusions

The results of the study confirm that teacher-made communicative English skills course examinations during 2008 to 2010 E.C. negatively influenced the teaching-learning process of the course. The analyses of the documents show that the syllabus of the course was designed based on the principles of communicative skills, but teacher-made communicative English skills course examinations were based on grammar-rules and reading comprehensions. Moreover, EFL tteachers used to give unfair emphases to some skills of the teaching materials during the construction of the examination items. Evidently, 61% (34% and 31 %) of the total items of the teacher-made examinations in all the tree years were reading and grammar. This led communicative English language EFL teachers of Ambo University to focus more on reading and grammar during teaching-learning processes. Other language skills such as speaking was given less emphasis while listening and writing skill were totally skipped which resulted in washback effect. The formats and the contents of all the three years examination were similar because the teachers of the course were relying on the previous examinations regardless of the table of specifications.

The examinations did not possess the fundamental qualities required in good language examination of all kinds. As a result, incompatibility between the examinations' items and objectives

to be assessed was apparent. The degree of the extent of relationship between the content area of the components of teaching materials and the teacher-made examinations was found to be very low. The correlation contingency coefficient between the two observations reads 0.38, 0.158, and 0.273 for those years. Moreover, cumulative contingency coefficient of the years reads 0.115 which shows very low relationship or harmful/negative washback effect of the teacher-made examination on the teaching–learning processes of the course. Negative washback effects of teacher-made examinations also influenced teachers' choice materials for teaching the course. It negatively affected student's confidence, attitude and perception towards the examination as a result students developed, stress, fear and anxiety towards examination; they usually come late to class and ignore the class the examination of the course.

Recommendations

The incomprehensive and disproportional distributions of the contents among the language contents within and between the teacher-made and teaching materials negatively affected the processes of the teaching learning of the course. Hence, embarking on the harmful effect of examinations, the teachers should first carefully draw up table of test specifications and develop a guideline to design the examinations in a very clear and precise items based on the particular language objectives, skills and component of the examinations. This test specification and guidance help the teachers to construct comprehensive and proportional examinations to the objectives of the teaching materials. For this end, EFL teachers should be given trainings to apply all the necessary test development stages including the Bloom's Taxonomy of instructional domains. The teachers should revise tests several times to avoid ambiguity and double answers for a given examination item. Ambo University should minimize the class size for English language teaching, prepare language laboratory and provide materials so the teachers can test listening skill.

References

Anderson, C. J. & Bachman F. L. (Ed.). 2006. *Assessing Young Language Learners*. Cambridge, New York: Cambridge University Press.

Ambachew M. D. 2003. *A Review of the Ethiopian Related Literature on Reading Ability*. University of Pretoria ltd. Pp. 83-93

Bailey, D. (1999). *'Washback in Language Testing. TOEFL Monograph Series'. Educational testing service.* TOEFL Program Office: Educational Testing Service.

Brown, D. 2004. *Language Assessment* . San Francisco State University : Longman.

Brown, J. D. 2012. Choosing the right type of assessment. In C. Coombe, S. J. Stoynoff, P. Davidson, & B. O'Sullivan (Eds.), The Cambridge guide to language assessment (pp. 133-139). Cambridge, UK: Cambridge University Press.

Burger, M, 2008. The Alignment of Teaching, Learning and Assessment in English language in grade 10 in district 9, Johannesburg (master thesis), University of South Africa

Cheng, L., Watanabe, Y., & Curtis, A. (Eds). (2004). *Washback In Language Tteting: Research Contexts and Methods.* London: LAWRENCE ERLBAUM ASSOCIATES, Publishers.

Ching, P. 2008. A critical review of five language washback studies from 1995-2007: Methodological considerations. *JALT Testing & Evaluation SIG Newsletter. 12 (2)*, 2-16.

Cohen, L. Manion, L. & Morrison, K. 2007. *Research Methods in Education (6th ed.).* Oxon: Routledge.

Creswell, J. W. (2009). *Research Design (4th ed.).* University of Nebraska-Lincon: SAGE Publications, Inc.

Davies, A. (2014). Fifty years of language assessment, In Antony Kunnan (ed.), *The Companion to Language Assessment*, 11(1) 1-18, John Wiley & Sons Inc, University of Edinburgh, Scotland, Retrieved on 28-Oct-2017 from http://onlinelibrary.wiley.com.

Djurić, M. 2008. Dealing with Situations of Positive and Negative Washback. *Scripta Manent 4(1)* , 14-27.

Green, A. 20007. Washback to learning outcomes: a comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education*, 75-97.

Gumaa, A. 2010. *The Impact of Test Content Validity on Language Teaching and Learning.* Shaqra University: Shaqra University press.

Hailu, D. 2015. *Assessing the content validity of teacher-made test in ELF class: the case of grade 11 at ambo preparatory (MA thesis).* Ambo University: Unpublished.

Hawkey, R. 2006. *Impact Theory and Practice.* New York: Cambridge University Press.

Hughes, A. 2003. *Testing for Language Teachers (2nd ed.).* Cambridge: Cambridge University Press.

Iyer, S. 2015. Impact of Washback in English as Second Language Classrooms – An Investigation. *ELT Voices- International Journal for Teachers of English Volume (5), Issue (3)*, 88-91.

Khan, I. S. 2013. Examination Washback Over Learning And Teaching. *Journal of Education and Practice*, 4-5.

Kothari. 2004. *Research Methodology (2nd ed.).* University of Rajasthan (India): New Age International Publisher .

Kumar, R. 2011. *Research Methodogy (3rd ed.).* Los Angeles: SAGE Publications.

Madsen, H. S. 1983. *Techniques In Testing.* Oxford University: Oxford University Press.

Messick, S. 1996. Validity and Washback in Language Testing. *Research Report*, 4-8.

Motuma, H. 2017. Alignmet of Primery school complition examinition with the textbook:2006-2009 infocuse. *Unpublished*.

Motuma, H. 2015.The Changing Role of the University TEFL Teacher in Improving the Students Communicative English Skill: From Passive Technician to Reflective Practitioner; Journal of Science and Sustainable Development (JSSD), 2015, 3(1), 133-155.

Motuma, H. 2014. An assessment of the Students' Reading Ability in Higher Institutions: The Case of Asela College of Teacher Education.*Journal of Science and Sustainable Development (JSSD)*, (1). 2. Pp.76-95.

Newfield, T. 2005. 'TOEIC Washback on Teachers. *A Pilot Study at One University*, 83-106.

Nugussie, T. 2002. Content Validity of the Ethiopian General Secondary Education Certificate English Examination (EGSECEE). *MA thesis* .

Pan, Y. 2009. *Do Tests Always Work?". The International Journal of Educational and psychological assessment.*

Sadeghi, S. 2014. High-stake Test Preparation Courses: Washback in Accountability Contexts. *Journal of Education & Human Development Vol. 3, No. 1*, 17-26.

Shih, C. 2007. The Canadian Modern Language Review. *A New Washback Model of Students' Learning'. The Canadian Modern*, 135-161.

Shannon-Baker, P. (2016). Making Paradigms Meaningful in Mixed Methods Research *Journal of Mixed Methods Research*, 10(4) 319–334.

Simachew, G. 2012. Washback of the University Entrance English Exam (UEEE) on Teachers' and Students' Practices: The Case of Preparatory Schools in Amhara National Regional State. *PHD Desertation*, 34-40.

Spratt, M. 2005. The implications for teaching and learning of studies of washback from exams. *Washback and the classroom*, 5- 29.

Sukyadi, D. & Mardiani, R. (2011). The Washback Effect of the English National Examination (ENE) on English Teachers' Classroom Teaching and Students' Learning. *VOLUME 13*, 96-111.

Taylor, L. 2005. Washback and impact. *key concepts in elt*, 154 .

Thaidan, R. 2015. Washback in language testing. *Washback in Language Testing. Education Journal*, 7.

Wall, D. 2000. The impact of high-stakes testing on teaching and learning: can this be predicted or

    controlled? *System, 28*, 499-509.

Weir, C. J. 2005. *Language Testing and Validation.* Palgrave Macmillan: Main's Press,.

Yasin, H. 2014. *Washbach effet of grade ten national examination on teachrs teaching practice

    (Master's thesis).* Haromaya : Haromaya university.

Zucker, S. 2005. Horizontal and Vertical Alignment. *Policy report*, 2.